

Detekování lidských tváří v obraze

Human Face Detection in Images

Zadání bakalářské práce

Student: **Ondřej Sahánek**

Studijní program: B2647 Informační a komunikační technologie

Studijní obor: 2612R025 Informatika a výpočetní technika

Téma: **Detekování lidských tváří v obraze**
Human Face Detection in Images

Zásady pro vypracování:

Vzhledem k vysokému nárůstu kamerových systémů v posledních letech se rozšiřují i systémy na detekování celé řady zájmových objektů. Uvažujeme jako zájmový objekt lidskou tvář. Systémy detekující lidské tváře mohou být využity například z bezpečnostních důvodů na letištích a podobných lokalitách. Dalším uplatněním detekce tváří v obrazech může být jejich hojně využívání ve fotoaparátech a o dalších nasazeních se v poslední době hovoří i v souvislosti se sociálními sítěmi.

1. Seznamte se s pojmy souvisejícími s touto problematikou.
2. Popište problémy, které mohou při detekování tváří nastat a které snižují úspěšnost detekce.
3. Zaměřte se na moderní aplikace detektorů tváří v obrazech.
4. Po nastudování předchozích bodů naimplementujte jednoduchý detektor. Využít můžete funkce z knihovny OpenCV.
5. V experimentální části proveďte zhodnocení a testování naimplementovaného detektoru.

Seznam doporučené odborné literatury:

Podle pokynů vedoucího diplomové práce.

Formální náležitosti a rozsah bakalářské práce stanoví pokyny pro vypracování zveřejněné na webových stránkách fakulty.

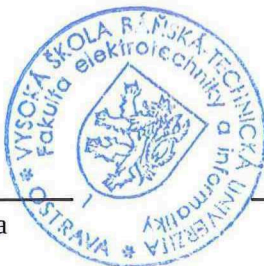
Vedoucí bakalářské práce: **Ing. Radovan Fusek**

Datum zadání: 18.11.2011

Datum odevzdání: 04.05.2012



doc. Dr. Ing. Eduard Sojka
vedoucí katedry



prof. RNDr. Václav Snášel, CSc.
děkan fakulty

Souhlasím se zveřejněním této bakalářské práce dle požadavků čl. 26, odst. 9 *Studijního a zkušebního řádu pro studium v bakalářských programech VŠB-TU Ostrava*.

V Ostravě 3.5.2012

Salva nek

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

V Ostravě 3.5.2012

Salva nek

Rád bych na tomto místě poděkoval mému vedoucímu Ing. Radovanu Fuskovi za velkou ochotu, trpělivost a cenné rady, které mi poskytl. Dále bych rád poděkoval rodině, za podporu, při vypracování mé bakalářské práce.

Abstrakt

Práce se zabývá problematikou detekce tváří. V teoretické části jsou popsány jednotlivé metody pro detekci a rozpoznávání obličejů. Pozornost je věnována metodám a technikám, používaných při detekci tváří, a faktorům ovlivňujících detekci. Cílem je získané informace použít při implementaci detektoru, za pomoci zvolené metody.

Klíčová slova: detekce tváří, metody

Abstract

The work deals with the problem of face detection. In the theoretical section describes various methods for detection and face recognition. Attention is paid to methods and techniques, used in the detection of faces, and factors affecting detection. The aim is to use gained information in the implementation of the detector, using the chosen method.

Keywords: face detection, methods

Seznam použitých zkratk a symbolů

RGB	– Barevný prostor složený ze tří složek, Red(červená), Green(zelená), Blue(modrá).
YCrCb	– Barevný prostor, kde Y zastupuje jas, Cr červenou barvu a Cb barvu modrou.
HSI	– Barevný prostor, kde H reprezentuje odstín, S sytost a I intenzitu.
GWN	– Gabor wavelet networks je metoda používaná k reprezentaci objektů.
PCA	– Principal component analysis, postup, pro redukci prostoru.

Obsah

1	Úvod	5
2	Rozpoznávání tváří	7
2.1	Detekce tváří	7
2.2	Extrakce příznaků	7
2.3	Rozpoznání obličejů	8
3	Problémy a faktory ovlivňující detekci a rozpoznání lidských tváří	9
3.1	Znalost lidských tváří	10
3.2	Faktory ovlivňující lidský vzhled	10
4	Metody rozpoznávání obličejů	13
4.1	top-down	13
4.2	bottom-up	13
4.3	texture-based	13
4.4	neural network	13
4.5	color-based	13
4.6	motion-based	14
4.7	depth-based	14
5	Detekce a rozpoznání obličejů založené na metodě Motion-based	15
5.1	Metoda rozpoznání obličejů ve video sekvencích a použití GWN	16
5.2	detekce a sledování	16
5.3	Extrakce a zpracování vlastních rysů	16
5.4	GWN, Gabor wavelet networks	18
5.5	Knn, K-nearest neighbor	20
6	Neuronové sítě	21
6.1	Neuron	21
6.2	Neuronové sítě	22
7	Color-based	25
7.1	RGB	25
7.2	YCbCr	26
7.3	HSI	27
8	Implementace detektoru lidských tváří	29
8.1	Volba barevného prostoru	29
8.2	Rozdělení složek barevného prostoru	29
8.3	Aplikování barevného modelu lidské kůže	30
8.4	Eliminace nežádoucích regionů	31
8.5	Označení	34

8.6	Výsledky	35
9	Závěr	37
10	Reference	39
11	Přílohy	41
11.1	Ukázky detekce obličejů na testovací sadě.	41

Seznam obrázků

1	Ukázka detekce obličejů na Facebook.com.	6
2	Detekce obličejů na digitalním fotoaparátu	6
3	Metoda rozpoznání obličejů.	7
4	Příklad celé procedury na vstupním obrázku. (a) Vstupní obrázek s detekovaným obličejem. (b) extrahovaný obličej. (c) vektor pevné velikosti po extrakci příznaků. (d) porovnání s databází a určení nejpravděpodobnější třídy.	8
5	Identická dvojčata představují značný problém, při identifikaci osob na digitální fotografii.	9
6	Software Olaface detekuje jak pohlaví tak i do jisté míry odhadne věk osob na fotografii.	9
7	Obličejové výrazy	11
8	Chybně označené tváře, díky matoucímu prostředí	11
9	Ukázka složení prostoru hlavních komponent	17
10	Úkázka reprezentace obličejových rysů za pomoci GWN.	19
11	K-nearest-neighbour.	20
12	McCulloch-Pittsův perceptron.	21
13	Dopředná síť.	22
14	Hopfieldova síť.	23
15	Reprezentace RGB barevného prostoru.	26
16	Lidský obličej převeden do barevného prostoru RGB.	26
17	Lidský obličej převeden do barevného prostoru YCbCr.	27
18	Reprezentace modelu barevného modelu HSI	28
19	Vstupní obraz pro algoritmus.	30
20	Jasová složka.	30
21	Složení červené barvy.	30
22	Složení modré barvy.	30
23	Obraz po aplikaci modelu pro detekci lidské kůže.	31
24	Výsledný obraz po aplikování redukce nežádoucích regionů.	32
25	Vstupní obraz.	33
26	Obraz po aplikování modelu lidské kůže.	33
27	Obraz po aplikaci redukce nežádoucích regionů.	34
28	Detekovaný obličej.	34
29	Mylná detekce.	35
30	Vliv pozadí a osvětlení.	42
31	Neúspěšná detekce.	42
32	Neúspěšná detekce.	43
33	Neúspěšná detekce.	43
34	Úspěšná detekce s regiony navíc.	43
35	Úspěšná detekce s regiony navíc.	44

1 Úvod

“The fantastic advances in the field of electronic communication constitute a greater danger to the privacy of the individual”

Earl Warren

Jak prohlásil Earl Warren, úžasný postup na poli elektronické komunikace představuje velké riziko pro soukromí jedinců. Není tomu jinak ani v oblasti detekce a rozpoznávání osob. Jakákoliv vaše fotografie, nebo video záznam, se stává prostředkem k vašemu monitorování. Stačí vás mít uložené v databázi, patřičný software, lokalitu, kde byla fotografie, či video pořízeno a tzv. time stamp (časovou známku), kdy byl záznam pořízen. V takovém případě ztrácíte veškeré soukromí, protože se lidé o vás mohou dozvědět vše, kde jste byli, s kým a taktéž kdy. Bohužel jistému omezení lidského soukromí se v dnešní době dobrovolně vzdáváme, neboť detekční a rozpoznávací systémy jsou běžné například na sociálních sítích. Využívají se k identifikaci osob jak v oboru kriminalistiky, tak i v oblasti bezpečnostních a přístupových systémů do objektů. Moderní prostředky pro záznam fotografií, či videozáznamů, již obsahují taktéž prvky pro detekci obličejových rysů.

V kriminalistice se nejčastěji využívá biometrických metod, forenzní portrétní identifikace, porovnání DNA, ale taktéž metody operativní a tipovací, kam patří sestavování portrétů zájmových osob, nebo porovnávání biologických materiálů.

V přístupu do zabezpečených objektů se často kombinují biometrické metody s přístupovým prvkem (čipová karta, klíč). Tyto metody jsou založeny například na kontrole váhy, analýze hlasu, nebo kontrole rohovky a duhovky.

Na sociálních sítích (Facebook) prvním krokem k rozpoznávání tváří bylo zprovoznění ”tagování” lidí na fotografiích. V červnu roku 2011 bylo zprovozněno automatické navrhování ”tagů” na fotografiích. Využívá k tomu systém izraelské firmy Face.com, která indexuje fotografie na Facebooku a přiřazuje k nim jména osob. Na jednu stránku tímto pomáhá lépe kategorizovat fotografie, ale na druhou stranu dochází k narušení soukromí, neboť lze lehce určit kdo, kde byl a s kým (obrázek 1).

Detekce obličejů v digitální fotoaparátch je v dnešní době běžná věc, ovšem s detekcí známých osob a určením osob, které fotíte nejčastěji, přišel jako první na světě Samsung.



Obrázek 1: Ukázka detekce obličejů na Facebook.com.

Detekce obličejů ve fotoaparátech funguje na principu detekování obrysů tváře. Fujifilm například využívá chip, který mezi stisknutím spouště a pořízením fotografie pořídí deset snímků, hledá na nich základní rysy jako oči, uši, nos, bradu. Výsledná tvář musí taktéž zaujímat minimálně 10 % fotografie. Poté zaostří a přizpůsobí expozici zájmové oblasti- obličej.



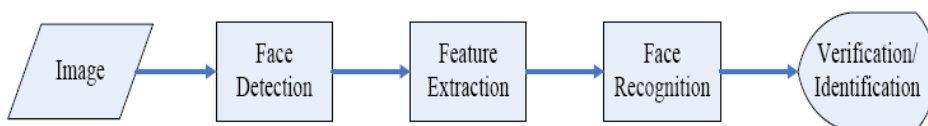
Obrázek 2: Detekce obličejů na digitalním fotoaparátu

Rozpoznávání obličejů je nejčastější a nejběžnější činností lidského podvědomí. Rozpoznávání lidských obličejů je velice komplexní záležitost ovlivněna mnoha faktory: jas, barvy, geometrie- natočení, přiblížení, emoční výrazy obličeje- smutek, smích, stárnutí, brýle, vousy. Je proto velice obtížné vytvořit detektor lidských obličejů, aby výsledek nebyl ovlivněn těmito rušivými vlivy.

2 Rozpoznávání tváří

Rozpoznávání tváří je metoda, která například na obrázku či videu určí, zda se obličej na snímku nachází. Pokud je tvář objevena, tak je lokalizována a nakonec určeno o čí tvář se jedná. Celý tento postup je rozdělen do tří základních kroků.

- Detekce tváří
- Extrakce příznaků
- Rozpoznání obličeje



Obrázek 3: Metoda rozpoznání obličejů.

2.1 Detekce tváří

Hlavním úkolem tohoto kroku je určit, zda se lidská tvář nachází na daném snímku. Pokud ano, tak určit i jeho polohu. Očekávaným výstupem tohoto kroku může být jak nový snímek s označenou lidskou tvář, tak i mnoho snímků s jednotlivými lidskými tvářemi. V případě výstupu mnoha různých snímků, je za účelem následného rozpoznání, jsou zaznamenány měřítka a orientace jednotlivých tváří. Kromě toho slouží i jako předzpracování snímku pro rozpoznávání tváří.

2.2 Extrakce příznaků

Hned po detekci lidských tváří, jsou tyto tváře extrahovány ze snímků. Přímé použití těchto tváří, pro rozpoznávání obličejů má jisté nevýhody. Za prvé většina těchto tváří se skládá z více jak 1000 pixelů, což je příliš mnoho pro rychlé a efektivní rozpoznání tváří. Za druhé, snímky mohou být pořízené z různých kamer, různých nastavení, s odlišnými výrazy ve tváři, rozdílným nasvícením.

Pro překonání těchto nedostatků se provádí extrakce příznaků, aby byly extrahovány pouze nezbytné informace, zmenšení rozměrů a čištění šumu. Zpravidla po tomto kroku je tvář změněna do vektoru pevné velikosti, nebo v sadu výchozích bodů a jejich odpovídající polohy.

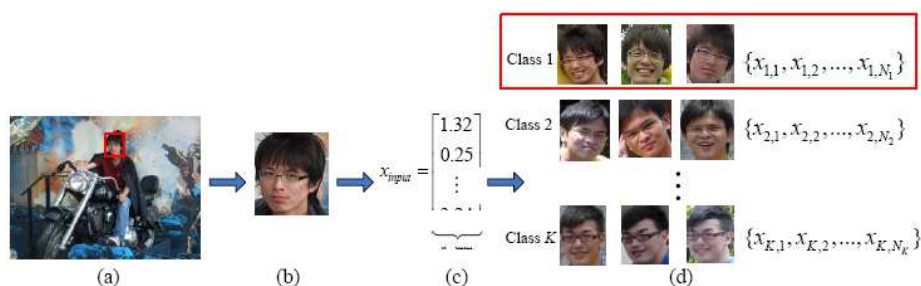
Někteří autoři uvádějí extrakci příznaků, jako součást detekce obličejů, či rozpoznání obličejů.

2.3 Rozpoznání obličejů

Po vygenerování zastoupení každé tváře, posledním krokem je rozpoznání identit jednotlivých tváří, neboli určení totožnosti. Za účelem automatického rozpoznávání tváří, je třeba vytvořit databázi tváří.

Pro každou osobu je pořízeno několik snímků, ze kterých jsou extrahovány rysy (příznaky) a následně uloženy v databázi. Poté, jakmile je pořízen vstupní snímek pro proces rozpoznání tváří, je na snímku detekován obličej a extrahovány příznaky. Následně jsou extrahované příznaky porovnány s těmi uloženými v databázi.

Existují dvě obecné aplikace rozpoznávání tváří, první je identifikace a druhá je nazývána ověření. Identifikace tváří znamená, že po systému chceme aby nám řekl o jakou konkrétní osobu se jedná, zatímco u ověření máme na vstupu snímek a odhad identifikace a systém určí zda je odhad správný, či nikoliv.



Obrázek 4: Příklad celé procedury na vstupním obrázku. (a) Vstupní obrázek s detekovaným obličejem. (b) extrahovaný obličej. (c) vektor pevné velikosti po extrakci příznaků. (d) porovnání s databází a určení nejpravděpodobnější třídy.

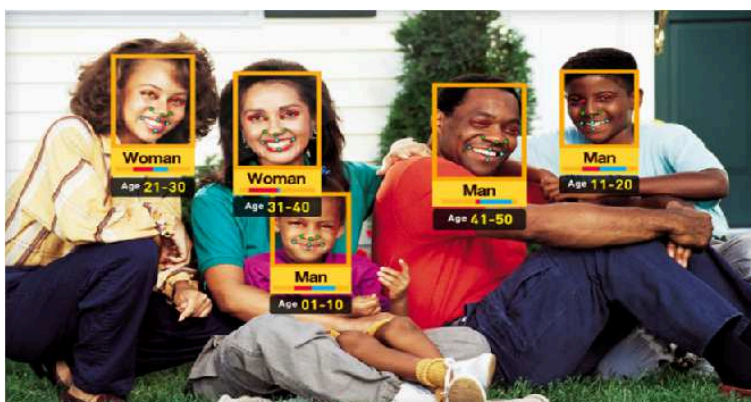
3 Problémy a faktory ovlivňující detekci a rozpoznání lidských tváří

Lidé se od počátku identifikují jako specifický jedinec a ne jen jako člověk. Každý z nás se snaží podvědomě identifikovat ne jenom osoby, které potkáváme, ale i zvířata, věci, zvukové a čichové podněty. Není tomu jinak ani v lidské společnosti. Všichni jsme lehce identifikovatelní pomocí občanských průkazů, otisků prstů, DNA, krevní skupiny, nebo v zabezpečených objektech pomocí oční duhovky. I když každý z nás je originální, v rozpoznání obličejů to ovšem nemusí být vždy tak zřetelně jasné. Například identická dvojčata představují značný problém pro rozpoznání a identifikaci (viz. obrázek 5).



Obrázek 5: Identická dvojčata představují značný problém, při identifikaci osob na digitální fotografii.

Stejně tak identifikace pohlaví představuje obrovský problém. Daleko v tomto oboru pokročila korejská společnost Olaworks, jejichž software "Olaface" je schopen detekovat jak osoby, ale tak i pohlaví a přibližný věk osob.



Obrázek 6: Software Olaface detekuje jak pohlaví tak i do jisté míry odhadne věk osob na fotografii.

3.1 Znalost lidských tváří

Existuje několik výzkumů v psychofyzice, zabývajících se problémem, jak lidé provádějí rozpoznávání. Mnoho z nich má přímé spojení k softwarovým programátorům zabývajících se návrhem algoritmů pro rozpoznávání tváří. Existují důkazy, že lidé si snadněji zapamatují tváře, než jakékoliv jiné objekty.

3.1.1 Holistický přístup vs. strukturální přístup

Holistický přístup tvrdí, že lidé rozpoznají tváře pomocí celkového vzhledu. Zatímco strukturální přístup je založen na domněnce, že identifikace tváří probíhá na základě obličejových rysů, jako jsou oči, nos, ústa, a také díky nim si obličej jedince zapamatujeme.

3.2 Faktory ovlivňující lidský vzhled

Existuje několik faktorů, které ovlivňují výsledek detektorů lidských tváří. Kromě nízké kvality snímku, natočení tváře, či okolního prostředí, ovlivněného procesem pořízení snímku, máme celkem šest faktorů, které musíme vzít v úvahu.

- osvětlení
- póza obličeje
- výraz obličeje
- RST (rotation, scale, translation) rotaci velikost a překlad
- matoucí prostředí
- zakrývání

3.2.1 Osvětlení

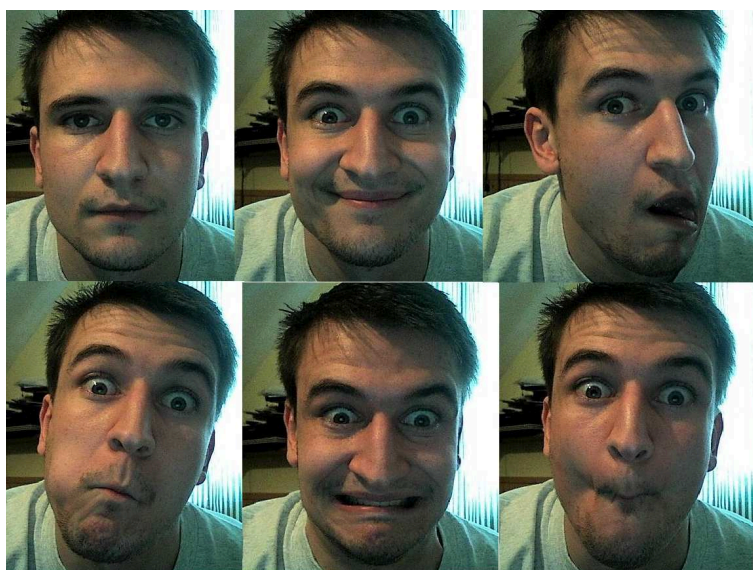
Tento faktor je způsoben, různým osvětlením prostředí. Změny osvětlení se zřetelně projevují na tvářích osob a v některých případech nemůžeme jednoznačně identifikovat osoby, v extrémních případech nedojde k detekci tváře samotné.

3.2.2 Póza obličeje

Póza obličeje vychází z různého úhlu a pozice během pořízení snímku. Tato vlastnost může změnit vzájemné vztahy obličejových rysů. Narušuje například metodu založenou na využití "eigenfaces".

3.2.3 Výraz obličeje

Lide používají různé obličejové výrazy, za účelem vyjádření citů a nálad. Emoční výrazy ne jenom, že vedou ke změně obličejových rysů, ale mohou i změnit tvar lidské tváře.



Obrázek 7: Obličejové výrazy

3.2.4 RST (rotation, scale, translation) rotaci velikost a překlad

Často vzniká při pořizování snímku. Vznikají problémy jak při detekci obličeje,

3.2.5 matoucí prostředí

K doplnění čtyř předešlých faktorů, musíme vzít v úvahu i pozadí a prostředí. Nevhodné pozadí může negativně ovlivnit výsledky detekce a poté i samotného rozpoznání.



Obrázek 8: Chybně označené tváře, díky matoucímu prostředí

3.2.6 Zakrývání

Znamená zakrývání obličejových částí pomocí, šátků, brýlí, vousů, čepic.

4 Metody rozpoznávání obličeje

V posledních letech, rozpoznávání obličejů připoutalo mnoho pozornosti a jeho výzkumu se nevěnují pouze v oboru informačních technologií, ale také například neurologové. Obzvláště detekce obličeje je důležitou součástí rozpoznávání obličejů jako první krok automatického rozpoznávání obličejů.

4.1 top-down

Předpokládá různé modely tváře na různých měřítkách v obraze. Pro úspěšnost je obrázek prohledáván v nejširším měřítku prvně. Jakmile je nalezena shoda, je obrázek prohledáván na dalším užším měřítku, dokud není dosaženo nejužšího možného měřítka. Obecně platí, že je předpokládán jeden model v každém měřítku (obvykle při čelním pohledu), je tedy velmi složité použít tuto metodu na různé pohledy.

4.2 bottom-up

Prohledává obrázek od "spodku-nahoru". Prohledává obrázek kvůli prvkům obličejových rysů, které jsou seskupeny do obličejových kandidátů, na základě jejich geometrických vztahů. Ačkoliv se dá tato metoda rozšířit na různé pohledy, je nemožné aby pracovala pod různými obrazovými podmínkami, protože obrazová struktura obličejových rysů se liší příliš mnoho, na to, aby se obličeje daly detekovat pomocí detektorů založených na detekci obličejových rysů.

4.3 texture-based

Tváře jsou detekovány zkoumáním prostorové distribuce šedých úrovní v obraze (SGLD matrix- Space Gray Level Dependency). Opět, tato metoda je špatně rozšiřitelná na různé úhly pohledů.

4.4 neural network

Detekuje tváře od podvzorků různých částí regionů obrazu, až po standardní velikost podobrazu a poté průchodem neuronovými sítěmi. Obecně algoritmus funguje velice dobře při čelním pohledu, ale výsledky se zhorší při použití algoritmu na jiné než čelní pohledy.

4.5 color-based

Funguje na metodě označení každého pixelu podle jeho podobnosti s barvou kůže. Následně každý region označí jako tvář, pokud obsahuje velkou koncentraci pixelů s barvou kůže. Je lehce rozšiřitelná na různé pohledy, problém ovšem nastává s citlivostí na barvu lidské kůže a tvar obličeje.

4.6 motion-based

Detekuje na základě odčítání pohyblivého popředí od statického pozadí. Tvář je poté lokalizována zkoumáním siluety, nebo pomocí barevných rozdílů obrazu. Metoda je nepoužitelná v případě, pokud je v zájmovém prostoru mnoho pohybujících se objektů.

4.7 depth-based

Hlavní obličejové rysy jsou detekovány pomocí informace založené na základě hloubky a intenzity. V prvním kroku je pár snímku z čelního pohledu odebráno z video sekvence. Poté jsou určeny rozdíly pomocí MHM¹ algoritmu. Na závěr jsou obličejové rysy určeny dle informace rozdílů v hloubce.

¹MHM-Multiresolution hierarchical matching algoritmus určení rozdílů v hloubce

5 Detekce a rozpoznání obličejů založené na metodě Motion-based

Metody, pro detekci obličejů, mohou být základně rozděleny na dva druhy. Statické a dynamické. Zatímco statické metody bývají jednodušší na implementaci a dosahují výborných výsledků, u dynamických metod tomu tak není.

Představme si situaci, kde máme detekovat všechny příchozí osoby do střeženého areálu v určitém časovém období. V takovém případě máme k dispozici video záznam z bezpečnostních kamer, bohužel na video záznamu se osoby pohybují, natáčejí, mění rysy tváře (díky emocím), mění se osvětlení v průběhu. Je zde mnoho faktorů které ovlivní úspěšnost samotné detekce a v neposlední řadě i samotného rozpoznání.

Rozpoznávání obličejů ve video sekvencích často zahrnuje tři důležité kroky:

- 1. Detekce obličeje
- 2. Sledování obličeje
- 3. Uznání

Detekce obličeje má za úkol lokalizovat oblasti, které mohou být označeny za obličej. Zatímco sledování obličeje, lokalizuje obličejové rysy například na základě polohy na předchozím snímku. Poslední uznání, identifikuje nebo ověřuje obličej.

Při detekci obličeje, jsou využity, nejčastěji metody založené na barevné informaci (color-based), pro segmentaci obličeje od pozadí snímku a k lokalizaci oblastí obličeje. Techniky pro detekci obličeje jsou velice podobné těm, které bývají použity u statických metod na fotografiích. Techniky sledování obličeje zahrnují sledování hlavy, kde hlava je brána jako pevný objekt, který mění polohu a rotaci, sledování obličejových rysů, kde deformace obličejových rysů, díky emocím, řeči, rotaci či pohybu, jsou viděny jako nepevné objekty a jejich transformace a rotace jsou omezeny anatomii hlavy. V neposlední řadě taktéž zahrnují sledování hlavy a rysů, tato technika bývá nejčastěji použita k rekonstrukci 3D modelu hlavy.

Hlavní problém pro detekci obličejů ve video sekvencích představuje nízká kvalita jednotlivých snímků ve video sekvenci. Přičemž nespornou výhodou je množství informací. To umožňuje výběr snímků, které budou použity pro detekci a rozpoznání obličeje a znovu použití informací získaných z předchozích snímků. Další nespornou výhodou je, že časová posloupnost umožňuje sledování obličejových rysů, díky tomu je do jisté míry eliminován negativní vliv emočních výrazů, zatímco pohyb, či chůze mohou zvýšit úspěšnost rozpoznání.

Je zřejmé, že obrovské množství dat obsažených ve videosekvenci, znemožňuje implementaci těchto čtyř kroků pro real-time software².

Aby bylo možné obejít problém s velkým množstvím dat, nejsou brány v úvahu všechny oblasti obličeje většinou systémů. Místo toho, jsou systémy navrženy tak, aby získával informace z rysových prvků obličeje, jako jsou oči, ústa či nos.

²Software poskytující odezvy okamžitě v aktuálním čase.

Jako ukázkou a popis této metody jsem zvolil postup trojice :Teófilo Emídio de Campos, Rogério Schmidt Feris a Roberta Marcondes Cesar Junior z Brazilské univerzity Instituto de Matemática e Estatística v Sao Paulu. Důvodem pro výběr tohoto postupu byl hlavně fakt, že na rozdíl od mnoha jiných postupů, bylo zde dosaženo vysoké úspěšnosti(97,7%) a to díky použití metody GWN³, a Knn⁴ algoritmu.

5.1 Metoda rozpoznání obličejů ve video sekvencích a použití GWN

Tento postup je rozdělen do čtyř klíčových bodů.

- 1. Detekce obličeje
- 2.Sledování obličeje
- 3.Extrakce obličejových rysů
- 4. Uznání

5.2 detekce a sledování

Detekce a sledování obličeje je důležitým krokem tohoto přístupu, neboť poskytuje normalizované obličejové rysy. Tento krok se dá dále rozdělit do dalších tří částí: detekce obličeje, detekce obličejových rysů a normalizace obličejových rysů.

Detekce obličeje se provádí za pomoci barevného modelu kůže a jednoduché procedury k ověření výskytu obličeje v každé detekované ploše označené za lidskou kůži. Jakmile je obličej detekován, je vypočítána jeho velikost a pozice.

Obličejové rysy, tj. oči, nos a ústa, jsou poté lokalizována a sledována za pomoci metody GWN. Tato metoda umožňuje detekci obličeje i v případě výskytu brýlí, vousů a různých emočních výrazů. Navíc metoda zohledňuje celkovou geometrii obličeje, díky tomu je imunní vůči situacím, kdy jedinec mrká, či se usmívá, které bývají kritické pro mnoho detektorů.

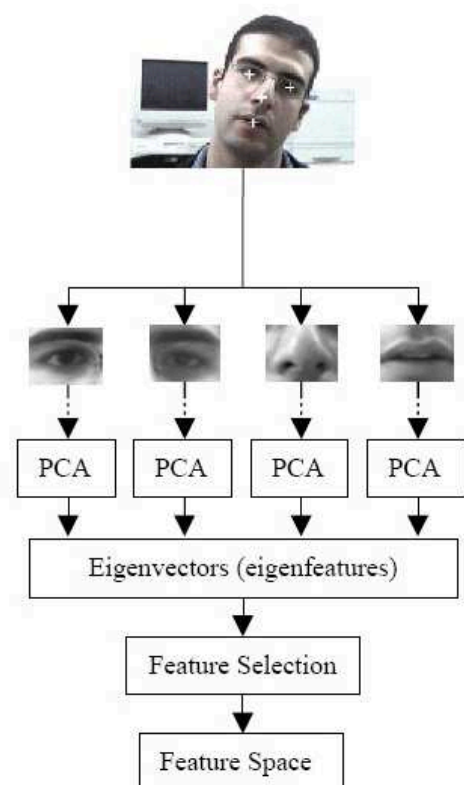
5.3 Extrakce a zpracování vlastních rysů

Po upřesnění pozice obličejových rysů je sada testovacích snímků ze sekvence použita k trénování čtyř základních obličejových rysů(pravé oko, levé oko, nos a ústa). Tyto objekty jsou dále použity k rozpoznání obličeje, neboť bylo dokázáno, že malé obrázky potřebují méně zpracování k naučení obličejových rysů , a rozpoznání obličejů. než kdyby se uchovával obličej jako celek. Za účelem rozpoznání tento postup generuje "feature space"⁵(obrázek 2.)

³Gabor wavelet networks- metoda pro reprezentaci objektů pomocí jednoduchých tvarů, určena k uchování informace o poloze, velikosti a natočení

⁴k-nearest neighbor

⁵jedná se o prostor složený z hlavních komponent (eigenfaces)



Obrázek 9: Ukázka složení prostoru hlavních komponent

Feature space je prostor složený z hlavních komponent, každá komponenta je reprezentována jako bod v n -rozměrném prostoru. Rozměr je určen počtem rysů použitých k popisu hlavní komponenty. Podobné komponenty bývají seskupeny dohromady.

5.3.1 Eigenfeatures(eigenfaces)

Po ustanovení polohy obličejových rysů a následné normalizaci oblastí výskytu obličejových rysů, je výsledný obraz použit k výcviku hlavních komponent⁶, za použití metody PCA⁷ známé též jako Karhunen-Loeve metoda pro redukci rozměru příznakového prostoru.

Každý eigenfaces představuje jen určité rysy obličeje, které mohou být zastoupeny v původním obrazu. Výhodou použití pouze obličejových rysů je výpočetní rychlost, protože menší obrázky vyžadují méně zpracování k PCA.

Eigenfaces, neboli hlavní komponenty vytvářejí velkou množinu. Představme si velkou videosekvenci, kde k detekci používáme určitý počet snímků. Z každého takového snímku mohou být vytaženy až 4 hlavní komponenty (oči, ústa, nos). I pro krátkou videosekvenci

⁶Eginefaces , obecně hlavní komponenty

⁷Principal component analysis

bude tato množina nabývat velkých rozměrů a né všechny hlavní komponenty přispějí k úspěšné detekci. Proto se za účelem dosažení větší úspěšnosti používají různé metody pro výběr hlavních komponent. V tomto zvoleném postupu se jedná o automatický výběr rysů.

5.3.2 Feature selection (výběr rysů)

Výběr rysů probíhá automaticky (automatic feature selection). Zjednodušeně řečeno, automatický výběr rysů je optimalizační technika, která vzhledem k množině rysů M , se pokouší určit podmnožinu rysů o velikosti D , kde $D < M$. To vede k maximalizaci některých kritérií, nezbytných, klíčových k následné detekci, a odstranění méně důležitých rysů, dle kterých by následná detekce, mohla být negativně ovlivněna.

Je dokázáno, že první tři eigenfaces snižují úspěšnost detekčních algoritmů, založených na detekci pomocí obličejových rysů. Toto je způsobeno tím, že tyto eigenfaces jsou ovlivněny změnou osvětlení a tím pádem i změnou barevného prostoru. Automatickým výběrem rysů je tomuto zamezeno, a to tak, že při výběru rysů potřebných k detekci, nejsou vybrány první eigenfaces z celé sady. Dosáhne se tím urychlení a hlavně lepších výsledků(zamezí se negativnímu vlivu při změně osvětlení).

5.3.3 Třídění

Posledním krokem nezbytným k úspěšnému rozpoznání lidských obličejů, byl v této metodě třídění jednotlivých obrazů. Za účelem seřazení a získání "feature space"⁸ byla použita metoda *Feature selection* založená na postupu, který vybírá eigenfaces tak, že vzdálenosti mezi jednotlivými prvky, které patří do různých tříd jsou maximalizovány, a pro prvky stejné třídy jsou minimalizovány.

Představme si prototyp jedné třídy jednoho snímku, například jistý snímek levého oka. Mějme množiny eigenfaces vybraných 10 snímků levého oka z 10 snímku videosekvence. Na každém snímku bude oko jinak natočené, nasvícené, může mít i různé rozměry. Prototyp každé třídy je určen jako hlavní vektor dané třídy. Třídící algoritmus tedy vybírá z dané třídy prvky nejbližší k prototypu, tím se zúží možnost výběru. Metoda je použita na každý jednotlivý hlavní prvek. Tento třídící algoritmus je nazýván *Knn*(K-nearest-neighbour, neboli K-nejbližšího-sousedů.).

5.4 GWN, Gabor wavelet networks

Gwn , neboli Gabor wavelet networks, je metoda používaná k reprezentaci různých objektů. Jedná se o kombinaci RBF⁹ networks a Gabor wavelets. GWN reprezentuje objekt jako lineární kombinace Gabor wavelets(dále jen GW), kde parametry jsou jednotlivými funkcemi GW(orientace, pozice a měřítko). GWN ve zpracování obrazu poskytují mnoho výhod

⁸množina vybraných eigenfaces

⁹Radial basis function network je umělá neuronová síť

- Do jisté míry jsou imunní proti změnám osvětlení a obrazovým deformacím.
- Gáborovy filtry jsou skvělými rysovými detektory.
- Přesnost reprezentace objektů může být volitelná, od naprosto autentické s fotografií, až po strohou reprezentaci pomocí bodů, či základních tvarů.

Použitím Gaborových filtrů je aplikován model k reprezentaci informací o objektech. GWN reprezentuje informace o objektech jako sadu primitivních obrázků, což vede ke značné redukci dat, potřebných ke zpracování obrazu. Informace o textuře, poloze a orientaci jsou zakódovány ve stejný čas, ale jsou rozděleny. Většina ostatních přístupů jako jsou PCA, nebo eigenfaces, berou v potaz většinou celé výřezy originálního obrazu, tzn. že neaplikují žádný model k reprezentaci objektů, místo toho reprezentace je na úrovni samotných pixelů. Z toho jasně vyplývá, že metoda GWN je časově i paměťově mnohem méně náročnější.

K definici GWN se dá použít souhrn funkcí. Vezmeme-li z množiny N podmnožinu lichých Gáborových funkcí

$$\gamma = (f_{n1} \dots f_{nN})$$

z rovnice

$$\begin{aligned} r_n(x+y) = \exp \left(-\frac{1}{2} \{ S_x((x-c_x)\cos\theta - (y-c_y)\sin\theta) \}^2 \right. \\ \left. + \{ S_y((x-c_x)\sin\theta + (y-c_y)\cos\theta) \}^2 \right) \\ * \sin(S_x((x-c_x)\cos\theta - (y-c_y)\sin\theta)) \end{aligned}$$

kde

$$n = (c_x, c_y, \theta, s_x, s_y)$$

zde c_x, c_y zastupují překlad, s_x, s_y zastupují zvětšení a θ určuje orientaci. Na následujícím obrázku je následná reprezentace obličejových rysů za pomoci metody GWN.



Obrázek 10: Ukázka reprezentace obličejových rysů za pomoci GWN.

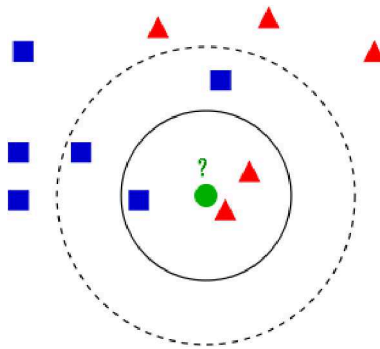
5.5 Knn, K-nearest neighbor

V rozpoznávání vzorů, je KNN metoda používaná ke klasifikaci objektů, na základě nejbližšího trénovacího prvku z "feature space". Knn je založeno na líném učení, kde jsou funkce pouze přiblíženy a samotné výpočty jsou provedeny až v klasifikaci. KNN je jeden z nejjednodušších algoritmů strojového učení. Objekt je ohodnocen většinovým hlasováním sousedů, s výsledkem, že je zařazen do třídy s nejvíce hlasy. K je kladné celé číslo, většinou nabývajících malých hodnot, pokud je K rovno jedné, potom je objekt jednoduše uložen do třídy nejbližších sousedů (1,2,3 atd...).

K výpočtu vzdálenosti se používá vzorec pro Euklidovskou vzdálenost:

$$d = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

kde q a p reprezentují jednotlivé body.



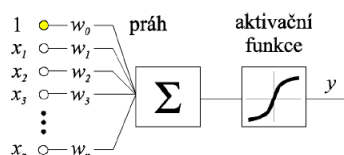
Obrázek 11: K-nearest-neighbour.

6 Neuronové sítě

Umělé neuronové sítě si jako svůj vzor berou Biologické neuronové sítě. Základem umělých neuronových sítí je neuron, který je taktéž inspirován svým biologickým vzorem. Výsledná neuronová síť je pak složena z mnoha takových uměle vytvořených neuronů. Jak už bylo zmíněno v metodách pro detekci obličejů, neuronové sítě se využívají pro detekci, ale taktéž pro kompresi dat, různé predikce, nebo pro klasifikace.

6.1 Neuron

Jako základní stavební jednotka neuronových sítí se používá neuron. Nejstarší a nejčastěji používaný je McCulloch-Pittsův perceptron.



Obrázek 12: McCulloch-Pittsův perceptron.

Vzorem McCulloch-Pittsůva perceptronu je biologický neuron. Takový neuron se skládá ze tří částí: vstupní, výpočetní a výstupní. Vstupní část je složena z několika vstupů. Tyto vstupy bývají oceněny a na základě ocenění probíhá upřednostňování jednotlivých vstupů. Výpočetní část se skládá z výpočetní nelineární funkce, někdy nazývané aktivační, nebo transferní funkce, která má za úkol zpracovat vstupní údaje a předat je výstupní části. Výstupní část se skládá z jednoho výstupu.

Celý neuron má za úkol přijmout informace, ty vyhodnotit a výslednou získanou informaci zpětně navázat na nějaký vstupní uzel jiného neuronu. Na první pohled velice podobné informačnímu systému, ale paměť neuronu netvoří ucelený prvek, nýbrž je rozprostřena do jednotlivých vstupních uzlů. Hlavní logická jednotka neuronu je taktéž velice primitivní a bývá tvořena jedinou nelineární funkcí, při čemž každý neuron vlastní funkci jinou.

Celý proces funguje součtem ocenění jednotlivých uzlů a následným vynásobením koeficientem. Na tento výsledek je aplikována výpočetní funkce a výsledek je pomocí výstupu navázán na vstup jiného neuronu. Celý neuron obsahuje ještě jednu konstantu tzv. prahovou hodnotu, která má za úkol aktivovat receptor. Stane se tak, když suma překročí prahovou hodnotu.

Funkce neuronu lze popsat následovně:

$$y = F\left(\sum_{i=1}^n (x_i * w_i + Z)\right)$$

kde x reprezentuje hodnotu na vstupu, w je cena vstupu, Z prahová hodnota, F nelineární funkce a y je hodnota výstupu.

6.2 Neuronové sítě

Jelikož samostatný neuron by nebyl schopen smysluplné činnosti, díky jednoduchosti výpočetních funkcí, skládají se neurony do umělých neuronových sítí. Jak tyto neurony vzájemně propojit určují topologie. Většinou se jedná o vrstvení neuronů, kdy je každý neuron propojen s každým. Každá síť musí obsahovat vstupní a výstupní vrstvu, nebo skryté, kdy všechny vrstvy mohou být součástí pouze jediné vrstvy.

Učení sítě obstarávají algoritmy, které mají za úkol měnit ocenění jednotlivých uzlů. teorie těchto algoritmů vychází z Hebbova zákona, který pochází z roku 1949. Říká nám, že se posilují vazby, které u neuronu způsobují jeho aktivaci. U umělých neuronových sítí se dá tento zákon vyjádřit následovně :

$$w_{i+1} = w_i + y_i * x_i$$

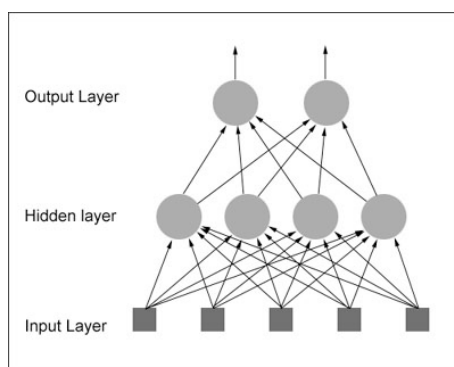
x_i je trénovací vektor hodnot vstupních atributů a y_i je informace o zařazení příkladu x do třídy. W_i reprezentuje ocenění před modifikací a W_{i+1} je ocenění po modifikaci. Je třeba důležité poznamenat, že i když se mění ocenění a následná priorita uzlů, na topologii sítě to nemá žádný vliv. Topologie sítě zůstává po celou dobu neměnná.

Některé topologie sítí:

- Dopředná síť
- Hopfieldova síť

6.2.1 Dopředná síť

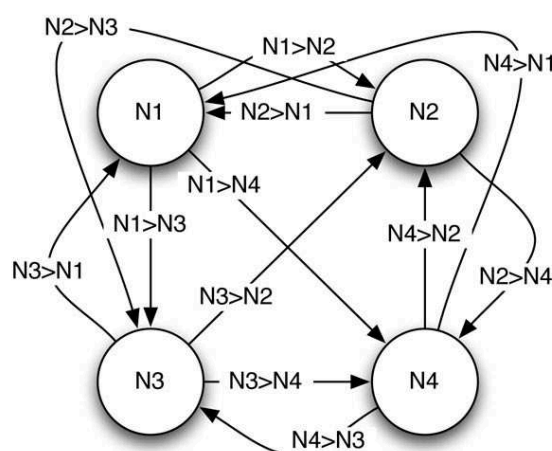
Je přímá technologie umělých neuronových sítí, obsahující skrytou vrstvu. Výstupy jsou nejdříve navázány na skrytou vrstvu a poté na výstupní. Jako učící algoritmus bývá často použit backpropagation. Algoritmus backpropagation zaručuje, že ihned po porovnání výsledného výstupu s očekávaným, má úprava ocenění v poslední (výstupní vrstvě) přednost před ostatními. Ihned poté následuje vrstva předešlá a tak sestupně až k vrstvě vstupní.



Obrázek 13: Dopředná síť.

6.2.2 Hopfieldova síť

Jedná se o jednovrstvovou síť vynalezenou Johnem Hopfieldem. Tato síť představuje model pro porozumění lidskému myšlení. Díky jednovrstevné architektuře, jsou všechny neurony zároveň součástí, vstupní i výstupní vrstvy. Všechny vstupy jsou oceněny. Tyto ocenění jsou měněné nezávisle na ostatních neuronech, v různý čas a u různých neuronů. Prahová hodnota, nebo aktivační hodnota, je binární. Na rozdíl od dopředné sítě, kde k aktivaci docházelo pouze v případě překročení prahové hodnoty, u Hopfieldovy sítě dochází k aktivaci v případě, že aktivační hodnota je nastavena na +1, druhá negativní hodnota je -1, v takovém případě je neuron neaktivní.



Obrázek 14: Hopfieldova síť.

Detekci obličejů za pomoci neuronových sítí zpracovali Stan Z. Li a Juwei Lu, kteří dosáhli výborných výsledků. Kdy v prvním kroku pomocí metody PCA redukovali příznakový prostor na testovací sadě obrázků 19 x 19 pixelů. V druhém kroku jako učící prostředek použili dvě identické neuronové sítě, vzájemně propojené, ale nezávislé, kde na vstupu u sítě A měli 22 prvků a na výstupu 12 a u sítě B na vstupu 12 a na výstupu 1. Síť B získává výstupy sítě A jako vstupy a výstup sítě B je finální výsledek. U obou sítí je použit učící algoritmus backpropagation. Neuronová síť A má za úkol odhadnout zda na obrázku se nachází obličej. Neuronová síť B má za úkol, v případě výskytu tváře, označit výstup za 1 a v případě výskytu "netváře" označit za -1. Jako "netvář" zde označují vše co nebylo právoplatně označeno za tvář, i kdyby se o tvář doopravdy jednalo.

Systém pracuje na odhadu pomocí zrcadlového převrácení obrazu. Jelikož zrcadlově obrácený obraz obličeje, se dá považovat taktéž za obličej, detekce na zrcadlovém obrazu by měla být taktéž úspěšná. Systém proto v případě pozitivní detekce obličeje, celý obraz zrcadlově obrátí a detekuje obličej i v tomto případě. Tímto se velice sníží pravděpodobnost omylů, za malou cenu výpočetního času.

Detekce za použití dvou neuronových sítí velice sníží počet tváří mylně označených, zároveň ovšem některé tváře neoznačí.

7 Color-based

Detekce obličejů založená na barvě, přináší velice dobré výsledky. Detekování obličejů za pomoci barvy lidské kůže přináší mnoho výhod. Image processing¹⁰ je mnohem rychlejší, než při zpracování ostatních obličejových rysů. Algoritmy potřebné k detekci jsou i jednodušší. Největší výhodou, kterou poskytuje tato metoda je, že barva lidské kůže je lehce reprezentovatelná v úzkém měřítku na barevném spektru. Tato reprezentace je stejná, pro lidi všech věkových kategorií, národností i obou pohlaví. To nesmírně usnadňuje práci a efektivitu algoritmů. Ovšem přináší to sebou i mnoho nevýhod. Detekce obličejů založená na barvě, je velice citlivá na nastavení barevných modelů, světelných podmínek. Detekce začíná být nepřesná v případě, že na zdrojové fotografii se vyskytuje více ploch označitelných jako lidská kůže, ovšem v profilových fotkách, kde se vyskytuje pouze obličej, na neutrálním pozadí, bez dalších ploch s lidskou kůží, je tato metoda rychlá přesná a spolehlivá. V případě, že na fotografiích se vyskytuje více obličejů, či barevná škála je pestřejší, nastává problém, jak určit které plochy jsou obličejem, a které nikoliv. Tento problém se dá řešit pouze za použití jiných postupů, jako jsou detekce hran, pomocí neuronových sítí, nebo pomocí určení obličejových rysů.

Algoritmy určené k detekci obličejů, založené na barevném modelu, musí jako první krok určit jaký barevný prostor bude využívat. Existuje jich celá řada:

- RGB (červená, zelená, modrá)
- YCbCr (jas, neboli osvětlení a modrá, červená)
- HSI(odstín, sytost, intenzita)
- CMY a CMYK(azurová, purpurová, žlutá, a černá)

Všechny prostory nesou různé výhody a nevýhody, v každém prostoru je barva lidské kůže jinak reprezentována. Některé prostory jsou více závislé na nastavení modelu, jiné na osvětlení, některé zase na intenzitě barev.

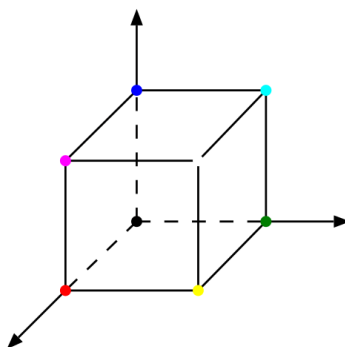
7.1 RGB

RGB je nejstarší a nejrozšířenější barevný model. Skládá se ze tří složek, red(červená), green(zelená), blue(modrá). Kombinací těchto tří složek v různém poměru vznikají barvy výsledné.

RGB model je reprezentován trojrozměrnou krychlí, s červenou, modrou a zelenou v rozích. Černá barva je na počátku a bílá na rohu protějším.

K číselné reprezentaci barev se používá 24 bitový systém s 8 bity vyhrazenými pro každou barvu. Červená je reprezentována pomocí [255,0,0], na trojrozměrné krychli [1,0,0]. RGB model zjednodušuje návrhy počítačových grafických systémů. Není bohužel ideální pro všechny aplikace. Jednotlivé barevné složky jsou navzájem velmi závislé a při změně

¹⁰Zpracování obrazu.



Obrázek 15: Reprezentace RGB barevného prostoru.

jedné dochází ke změně i ostatních dvou. Toto znemožňuje implementaci pro některé systémy, jako vyrovnaní histogramů.



Obrázek 16: Lidský obličej převeden do barevného prostoru RGB.

Využití RGB v detekci tváří najde své uplatnění, neboť lze jednoduše detekovat pixely, jejichž bitové hodnoty barev, jsou v rozmezí barevného modelu lidské kůže. Tyto pixely, poté můžou být upraveny pro případné změny v osvětlení, či intenzitě barvy. Za účelem detekce obličejů se využívají většinou složky R a G, pomocí kterých lze lehce formulovat barevný model lidské kůže. Jakmile je lidská kůže, detekována je označena za tvář. Tento postup ovšem selhává, pokud je na fotografii více oblastí s lidskou kůží.

7.2 YCbCr

Barevný prostor YCbCr byl definován, díky narůstajícím požadavkům na algoritmy, zpracovávající videosekvence. Patří do rodiny barevných prostorů, využívajících se v televizním přenosu. Skládá se ze tří prvků, osvětlení (jasové složky), barvy modré a barvy červené. K číselné reprezentaci je opět použit 8 bitový systém, kde pro $Y=16$ je repre-

zentován prostor černé a pro $y=235$ je reprezentován prostor bíle. Barevná složka červená a modrá jsou reprezentovány v rozsahu 16 - 240.



Obrázek 17: Lidský obličej převeden do barevného prostoru YCbCr.

K implementaci modelu lidské kůže, ve většině algoritmů postačí barevné složky Cb a Cr. Ovšem k přesnějšímu určení modelu a tedy i přesnější detekci se využívá i jasová složka Y. Nese to ovšem sebou riziko nefunkčnosti detekce, při velkých rozdílech v osvětlení. Algoritmus založený na tomto modelu, by měl taktéž obsahovat vstupní data pouze s lidskými obličejí, ze stejných důvodů jako u modelu RGB.

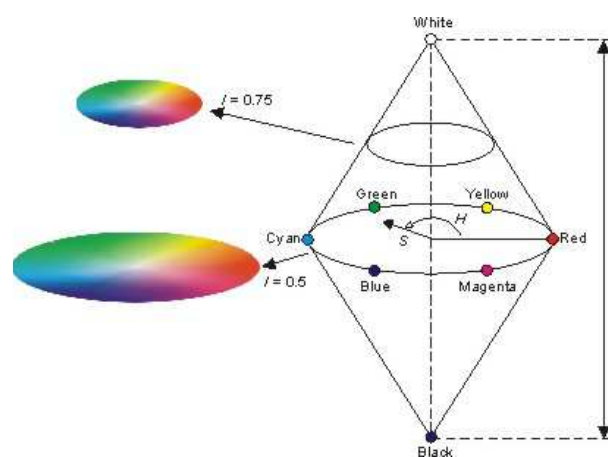
7.3 HSI

Vzhledem k tomu, že odstín, sytost a intenzita jsou tři vlastnosti používané k popisu barvy. Je logické, že vznikl odpovídající barevný model, HSI. K reprezentaci barev nemusíte znát jednotlivý poměr. Žluté, nebo modré barvy. Jednoduše stačí přizpůsobit odstín barvě, kterou chcete získat. Pokud chcete mít z červené růžovou, změníte sytost. Pokud chcete mít zase barvu tmavší, změníte intenzitu. Tento barevný prostor je často využíván pro operace spojené s Histogramy.

K reprezentaci HSI modelu se využívá dvojitého kuželu. Odstín je reprezentován jako úhel. Sytost odpovídá poloměru a intenzita se liší podle svislé osy od 1 (bílá) po 0 (černá).

K reprezentaci lidské kůže v barevném modelu HSI jsou použity složky odstínu a sytosti. Stejně jako oba dva předešlé barevné modely, i tento má problém s detekcí obličejů v případě výskytu více ploch shodných s modelem lidské kůže.

Jak je patrné u metody detekcí obličejů je jedna velká nevýhoda. I když, jak je všude uváděno, že detekce dle barevného modelu lidské kůže je velice spolehlivá, opak je pravdou. Velice spolehlivá je v případě, že máme fotografii či video-záznam s ničím nekrytým obličejem, s neutrálním pozadím, barevně odlišným od barvy lidské kůže.



Obrázek 18: Reprezentace modelu barevného modelu HSI

8 Implementace detektoru lidských tváří

K implementaci detektoru, byla zvolena metoda založená na barevném modelu lidské kůže. Detektor je naprogramován za pomoci knihovny OpenCV.

OpenCV je volně dostupná knihovna, určená primárně pro práci s obrazy a video-sekvencemi. Obsahuje mnoho implementovaných funkcí. Například pro detekci hran, detekci Haarových příznaků a mnoho dalších. Je kompatibilní s C, C++, Python, nebo Java, spustitelný jak pod Windows, Linux, Android a Mac. Jako základní postup pro detekci obličejů, jsem použil metodu popsanou v[10]. Jay Kapur zde uvádí základní metodu založenou na barvě.

8.1 Volba barevného prostoru

Jako barevný prostor byl nejdříve zvolen barevný prostor RGB. Tato volba se ukázala být mylnou, neboť po implementaci celého algoritmu, byly výsledky velice neuspokojivé. Na testovacím vzorku 40 obličejů různých lidí, různých národností a rozdílných pohlaví, dosahoval algoritmus úspěšnosti 32,5 % (13 úspěšných detekcí ze 40 vstupních vzorků). Z toho u všech třinácti detekovaných obličejů byly i jako obličeje mylně označeny i prostory mimo tvář.

Na vině byl barevný prostor, který u RGB (viz. obr. 16) reprezentuje model lidské kůže, jako necelistvý prostor. Je velice citlivý na změny intenzity v závislosti na osvětlení. To vedlo k rozčlenění obličeje na mnoho úseků a vzniku necelistvých ploch.

Tento problém byl vyřešen změnou barevného modelu a to YCbCr. Tento barevný model, vytváří pro detekci obličejů mnohem lepší podmínky. Díky jasové složce a jisté toleranci v této složce, vytváří celistvý prostor, lidské kůže (viz. obr. 17.). To umožňuje pracovat s lidskou kůží jako jediným prostorem a nesmírně usnadňuje detekci.

K překladi do barevného prostoru YCbCr je využíván následující vzorec.

$$Y = +0,299 * R + 0,587 * G + 0,114 * B$$

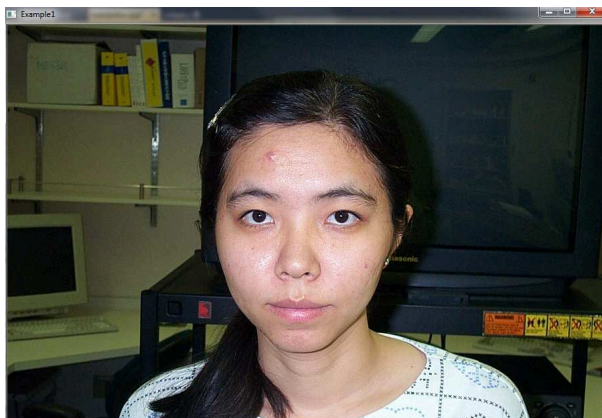
$$Cb = -0,169 * R - 0,331 * G + 0,5 * B$$

$$Cr = +0,5 * R - 0,419 * G - 0,0813 * B$$

Kde R je reprezentace červené barvy v RGB, G je reprezentace zelené barvy a B barvy modré.

8.2 Rozdělení složek barevného prostoru

Dalším nezbytným krokem je rozčlenění barevného prostoru na jednotlivé složky. Obraz převedený do barevného prostoru YCbCr je na obrázku 17. Jednotlivé složky jsou dále extrahovány a rozděleny na samostatné prvky. OpenCV k tomu využívá funkci cvSplit. Jak je zřejmé na kanálu červené i modré barvy začíná být zřetelná lidská kůže. Za lidskou kůží, lze ale označit i některé objekty na pozadí.



Obrázek 19: Vstupní obraz pro algoritmus.



Obrázek 20: Jasová složka.



Obrázek 21: Složení červené barvy.



Obrázek 22: Složení modré barvy.

8.3 Aplikování barevného modelu lidské kůže

Dalším nezbytným krokem je aplikace modelu lidské kůže. Mnohokrát už bylo zmíněno, že lidská kůže se dá definovat pomocí úzké škály na barevném spektru. V [11] definovali model lidské kůže v barevném modelu YCbCr, následovně:

$$Y > 80$$

$$85 < Cb < 135$$

$$135 < Cr < 180$$

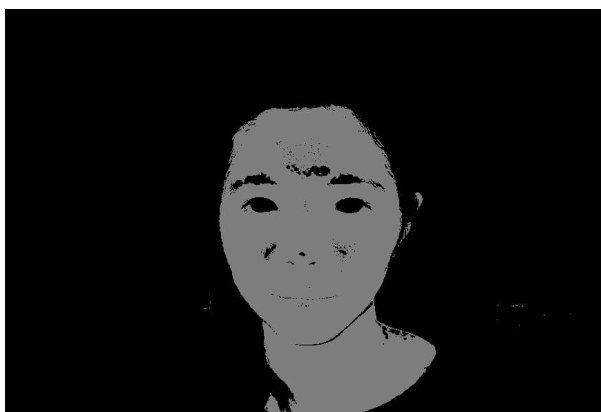
Většina modelů barvy lidské kůže je téměř identická, liší se pouze posunem škály nahoru či dolů o pár bodů, nebo zúžením prostoru, v případě zaručení některých podmínek, jako stálého osvětlení, stejný úhel snímání a podobně. Tento barevný model bohužel vykazoval neuspokojivé výsledky. V případě, že pozadí snímku bylo příliš světlé (pozadí a kůže se v podstatě sjednotila). Proto jsem barevný model upravil tak, aby k detekci docházelo i v případě výskytu světlého pozadí. Celý model lidské kůže jsem upravil následovně.

$$80 < Y < 200$$

$$100 < Cb < 130$$

$$140 < Cr < 170$$

Při extrakci lidské kůže dojde k největším změnám u barevné složky Cb. Dalo by se tedy říci, že detekce by byla možná pouze v závislosti na složce Cb, ale není tomu tak, neboť při opětovném složení jednotlivých složek modelu YCbCr, dojde k eliminaci nežádoucích prvků.



Obrázek 23: Obráz po aplikaci modelu pro detekci lidské kůže.

Jak je zřejmé po aplikaci modelu by se dalo téměř s jistotou určit, kde se nachází lidská tvář. Ovšem metoda pro označení lidské tváře bere v potaz všechny regiony. Tudíž by byly označeny i nežádoucí malé elementy na pravé a levé straně od tváře, zobrazené jako bílé (našedlé) místa. Toto je největší problém při detekci založené na barevném modelu lidské kůže. Takové místa bývají často označené za tváře, i když tomu tak ve skutečnosti není.

8.4 Eliminace nežádoucích regionů

Jak je popsáno, nevýhodou metody založené na aplikaci barevného modelu lidské kůže, je označení regionů, které obličej nejsou, právě za detekované obličej. I zde by v takovém případě bylo za tvář označeno 12 objektů.

V openCV pro detekci nežádoucích objektů a extrakci jen objektů našeho zájmu, existuje mnoho knihoven a funkcí například se dá využít knihovny cvBlob, metody cvHaarDetectObjects pro detekci Harových příznaků, nebo cvHOGDescriptor (nově obsažná knihovna ve verzi 2.4.0).

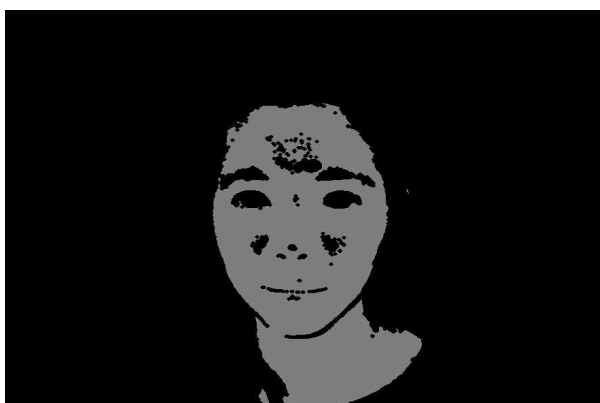
Jelikož po celou dobu implementace jsem se vyhýbal využívání cizích knihoven, jsem přemýšlel i nad způsobem jak odstranit nežádoucí objekty z obrazu tak, aby ve výsledku bylo k označení za obličej co nejméně ploch, v nejlepším případě pouze jedna.

Při neustálé kontrole výsledku, po aplikaci modelu kůže, jsem zkoumal obraz a nežádoucí regiony. Ve většině případů se jednalo o malé značky, dřevěné obložení, knihy, proužky na košili a další drobné objekty. Ve výsledku se vždy jednalo o malé nespojitě body, nebo obrysy. Tyto body, nebo obrysy se dají jednoduše detekovat a překreslit.

Aplikoval jsem tedy jednoduché pravidlo pro detekci obrysů (v openCV contours), které hledá nespojité obrysy, či body. Výhodou je, že body jsou brány jako nespojité obrysy. Tyto obrysy(body) byly poté překresleny na negativní barvu (černou). Obrys obličeje za každých okolností tvoří spojitý obrys, stejně tak ale i ruce, či nohy.

Pravidlo se dá popsat jednoduše: hledej nespojité křivky, když nalezneš nespojitou křivku, překresli ji čarou tloušťky 3 pixelů, do hloubky 5 pixelů. Rozumějte hloubky jako 5 sousedních pixelů od nalezeného bodu, nebo obrysu.

Výsledkem je zrušení velkého počtu nežádoucích regionů, za cenu, že v oblasti obočí, či vousů, dojde k odmazání pixelů s modelem kůže. V takovém případě to nemá žádný vliv na finální detekci, neboť v té se bere v potaz objekty jako celek s uzavřeným obrysem.



Obrázek 24: Výsledný obraz po aplikování redukce nežádoucích regionů.

Jak lze vidět na obrázku 24. došlo k smazání nežádoucích obrysů, či bodů. V tomto případě jich bylo 431.

Pro názornější ukázkou funkčnosti a efektivnosti tohoto pravidla, jsem jako vstupní obraz použil hromadnou fotografii 45 lidí, různého pohlaví, různých národností, různého věku, oblečení.



Obrázek 25: Vstupní obraz.



Obrázek 26: Obraz po aplikování modelu lidské kůže.

Jak jde vidět na rozdíl mezi obrázkou 26 a 27, došlo k velké redukci, přesněji bylo smazáno 7022 nežádoucích obrysů či bodů. Byly redukovány regiony jako jsou šaty, klobouky, stěny v pozadí. Za výsledek to má obrovské snížení mylné detekce.



Obrázek 27: Obraz po aplikaci redukce nežádoucích regionů.

8.5 Označení

Posledním krokem v detekci obličejů je zřetelné označení nalezených obličejů. Jedná se o samotnou detekci. Kdy všechny regiony s barvou lidské kůže, jsou označeny za tvář. Takový region je poté na výsledné fotografii nějakým způsobem označen, ohraničen, jako důkaz úspěšné, či neúspěšné detekce.



Obrázek 28: Detekovaný obličej.

Pro ukázkou mylné detekce, jsem použil již zmíněný snímek 45 osob.

Při bližším zkoumání si lze všimnout, že mnoho obličejů bylo označeno správně, ale poté velké množství regionů mylně, regiony jako ruce, krk, nohy, nebo díky spojitosti regionů, bylo měřítko příliš veliké a jako obličej je označena příliš velká plocha.



Obrázek 29: Mylná detekce.

8.6 Výsledky

Testování algoritmu jsem provedl na 40 snímcích z testovací databáze [13], různých lidí, různých národností, za různých světelných podmínek. Jednalo se o snímky tváří s různým pozadím, a minimálním počtem výskytů jiných částí lidského těla. Úspěšně detekovaných tváří bylo 82,5% (33). Z toho u 13 snímků bylo označeno několik regionů navíc jako obličej. Jedná se o části těla, jako jsou ruce, uši, krk. U 5 snímků byly označeno více regionů jako tváře. V tomto případě se jednalo o objekty na pozadí jako knihy, značky, plakáty.

Za pomoci barevného prostoru YCbCr, bylo dosaženo mnohem větší úspěšnosti, než za použití RGB, i mylná detekce byla redukována na výskyt u 54 % testovaných snímků (jen snímky u kterých byla správně detekována tvář). Kde 72% mylných detekcí je způsobeno regiony jako jsou uši, krk, ruce.

9 Závěr

Dle zadání se práce zabývá metodami, používanými pro detekci a rozpoznávání obličejů. V teoretické části jsou vybrané metody popsány. Je zde nastíněn postup pro detekci obličejů, za pomoci pohybu na video záznamech a prostředky používané při této metodě. Dále se práce zabývá metodou s využitím neuronových sítí a metodou založenou na barevném modelu lidské kůže.

Po nastudování jednotlivých metod, byla pro implementaci zvolena metoda založená na detekci barvy lidské kůže (color-based). Tato metoda byla zvolena z důvodu prohlašované vysoké úspěšnosti při použití. K implementaci byla použita doporučená knihovna OpenCV, jejíž využití byla umožněna díky zadání.

V průběhu implementace bylo zjištěno mnoho nevýhod, souvisejících se zvoleným barevným modelem RGB. Výsledkem byla nízká úspěšnost implementace. Byl zvolený nový postup, založený na barevném modelu YCbCr, který přináší mnohem lepší výsledky.

Dále práce odhaluje nedostatky zvolené metody a to v podobě mylné detekce, citlivosti na osvětlení a velké závislosti na nastavení, či závislosti na vstupních datech. Mylná detekce spočívá v označení regionů za tváře, i když se jedná pouze o části lidského těla, nebo v horším případě za cizí objekty.

V části Eliminace nežádoucích regionů je navržen postup, pro redukci těchto nedostatků, za účelem snížení mylné detekce. Tento jednoduchý postup dosahuje až překvapivě dobrých výsledků a zefektivnění celé implementace.

Do budoucna, by se dala metoda kombinovat s dalšími postupy, jako detekcí obličejových rysů, a tím zefektivnit detekci obličejů. Základní metoda detekce, založená na barevném modelu kůže, může být využita u většiny dalších metod, například motion-based, díky poskytnutí značné redukce obrazového prostoru ke zpracování. S detekovanou tváří by bylo možné i nadále pracovat, například extrahovat pro další zpracování, počítat výskyty tváří pro statistické účely.

Metoda založená na detekci barvy může být využita i pro jiné implementace, než jen detekce tváří v obrazech. Dá se využít například pro detekci automobilů, pomocí barvy světel, detekci dopravních značek, k detekci látek na mikroskopických snímcích.

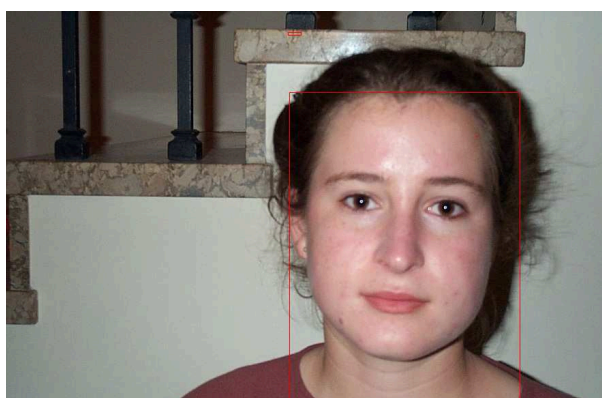
10 Reference

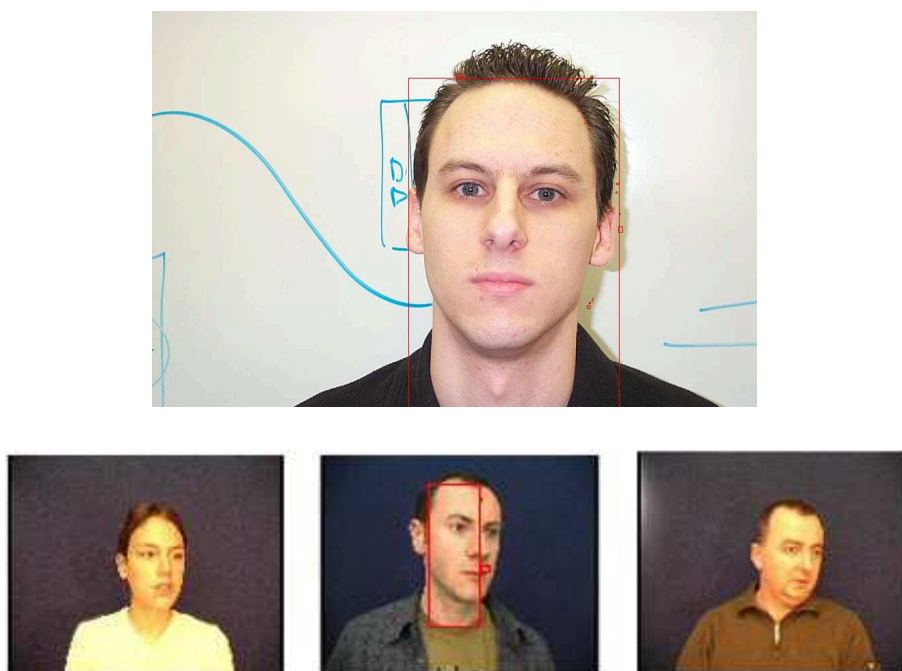
- [1] Sanjay Kr. Singh D.S. Chauhan, Mayank Vatsa, Richa Singh *A Robust Skin Color Based Face Detection Algorithm*, dostupné z :< [http : //www.csee.wvu.edu/ richas/papers/tkjse.pdf](http://www.csee.wvu.edu/richas/papers/tkjse.pdf) >
- [2] Patrick Zandl, *Rozpoznávání fotografií je další velká věc. Jen se neví, zda dobrá* dostupné z :< [http : //www.lupa.cz/clanky/rozpoznavani – fotografii – je – dalsi – velka – vec – jen – se – nevi – zda – dobra/](http://www.lupa.cz/clanky/rozpoznavani-fotografii-je-dalsi-velka-vec-jen-se-nevi-zda-dobra/) >
- [3] Teófilo Emídio de Campos, Rogério Schmidt Feris, Roberto Marcondes Cesar Junior *A Framework for Face Recognition from Video Sequences Using GWN and Eigenfeature Selection*, dostupné z :< [http : //rogerioferis.com/publications/FerisWAICV00.pdf](http://rogerioferis.com/publications/FerisWAICV00.pdf) >
- [4] Wei-Lun Chao *Face Recognition* GICE, National Taiwan University dostupné z : < [http : //www.lrv.fri.uni – lj.si/ peterp/publications/cvww99.pdf](http://www.lrv.fri.uni-lj.si/peterp/publications/cvww99.pdf) >
- [5] Jiří Hinner *Detekce a rozpoznávání obličejů osob a jejich identifikační význam* dostupné z : < [http : //aplikace.mvcr.cz/archiv2008/casopisy/kriminalistika/2003/03₀₁/hinner.html](http://aplikace.mvcr.cz/archiv2008/casopisy/kriminalistika/2003/03_01/hinner.html) >
- [6] Volker Krueger and Gerald Sommer *Gabor Wavelet Networks for Object Representation* dostupné z : < [http : //citeseerx.ist.psu.edu/viewdoc/download?doi = 10.1.1.79.4673rep = rep1type = pdf](http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.79.4673rep=rep1type=pdf) >
- [7] Peter Peer, Franc Solina *An Automatic Human Face Detection Method* dostupné z :< [http : //www.lrv.fri.uni – lj.si/ peterp/publications/cvww99.pdf](http://www.lrv.fri.uni-lj.si/peterp/publications/cvww99.pdf) >
- [8] Vladimir Veyhnevet, Vassili Sayonov, Alla Andreeva *A Survey on Pixel-Based Skin Color Detection Techniques* dostupné z :< [http : //www.graphicon.ru/oldgr/en/publications/text/gc2003vsa.pdf](http://www.graphicon.ru/oldgr/en/publications/text/gc2003vsa.pdf) >
- [9] Stan Y.Li Juwei Lu *Face Detection by Neural Learning* dostupné z :< [http : //www.graphicon.ru/oldgr/en/publications/text/gc2003vsa.pdf](http://www.graphicon.ru/oldgr/en/publications/text/gc2003vsa.pdf) >
- [10] Jay P. Kapur *Face detection in Color images* dostupné z :< [http : //web.archive.org/web/20090723024922/http : //geocities.com/jaykapur/face.html](http://web.archive.org/web/20090723024922/http://geocities.com/jaykapur/face.html) >
- [11] JORGE ALBERTO MARCIAL BASILIO, GUALBERTO AGUILAR TORRES, GABRIEL SÁNCHEZ PÉREZ, L. KARINA TOSCANO MEDINA, HÉCTOR M. PÉREZ MEANA *Explicit Image Detection using YCbCr Space Color Model as Skin Detection* dostupné z :< [http : //www.wseas.us/e – library/conferences/2011/Mexico/CEMATH/CEMATH – 20.pdf](http://www.wseas.us/e-library/conferences/2011/Mexico/CEMATH/CEMATH-20.pdf) >

- [12] Moritz Störing, Hans K. Andersen, Erik Granum *Skin colour detection under changing lighting conditions* dostupné z :< [http : //web.archive.org/web/20070412154255/http : /www.vision.auc.dk/ mst/Publications/sirs99.pdf](http://web.archive.org/web/20070412154255/http://www.vision.auc.dk/~mst/Publications/sirs99.pdf) >
- [13] Caltech faces database *databáze obličejů* dostupné z :< [http : //www.vision.caltech.edu/html – files/archive.html](http://www.vision.caltech.edu/html_files/archive.html) >
- [14] Robert Laganière *OpenCV 2 Computer Vision Application Programming Cookbook* Packt Publishing Ltd, 2011
- [15] Gary Bradski, Adrian Kaehler *Learning OpenCV* O'REILLY , 2008
- [16] <http://upload.wikimedia.org/wikipedia/commons/e/e7/KnnClassification.svg>

11 Přílohy

11.1 Ukázky detekce obličejů na testovací sadě.

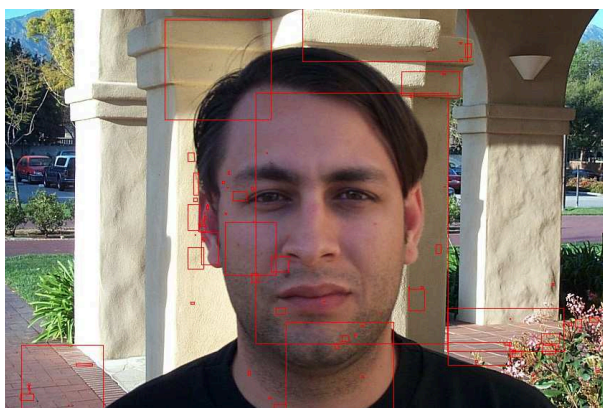




Obrázek 30: Vliv pozadí a osvětlení.



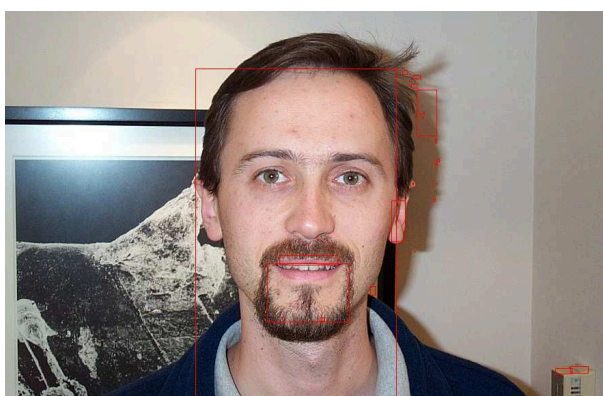
Obrázek 31: Neúspěšná detekce.



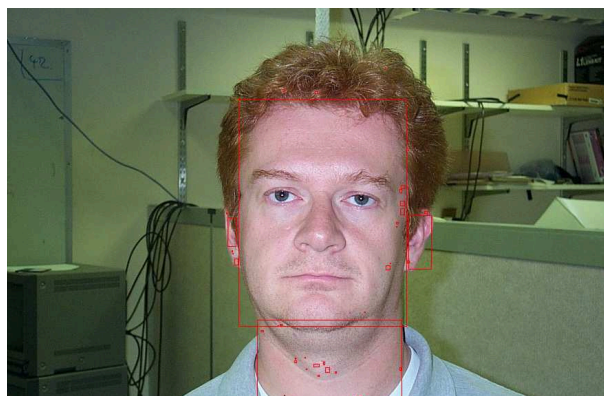
Obrázek 32: Neúspěšná detekce.



Obrázek 33: Neúspěšná detekce.



Obrázek 34: Úspěšná detekce s regiony navíc.



Obrázek 35: Úspěšná detekce s regiony navíc.